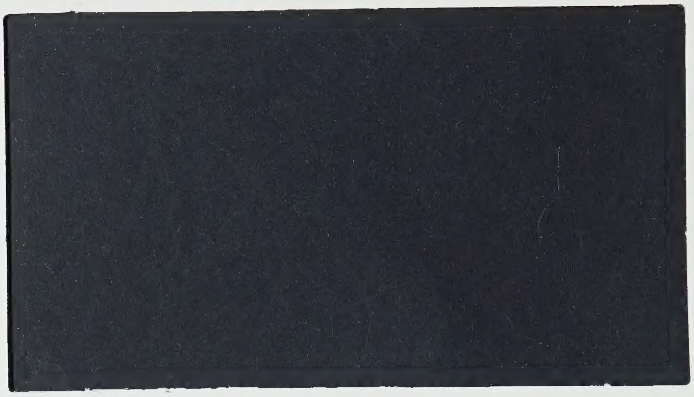
Canada. Commission of Inquiry
on the Pharmaceutical Industry
Issues in patent policy with
respect to the pharmaceutical
industry (background study)

Background Study

ISSUES IN PATENT POLICY WITH RESPECT TO THE PHARMACEUTICAL INDUSTRY

Commission of Inquiry on the Pharmaceutical Industry

Canada



ISSUES IN PATENT POLICY WITH RESPECT TO THE PHARMACEUTICAL INDUSTRY

Y. Kotowitz

Department of Economics
University of Toronto

Cat. No. RG48-1/6-1986E
ISBN 0-662-14497-4

CCAC No. 192 25166 E 86-05

Table of Contents

Y. Kotowitz.

EXECUTIVE SUMMARY

General Rationale for the Patent Systems

In order to undertake R&D firms must have prospective profits. It is well established that the volume of R&D is positively related to potential profitability. Serious appropriability obstacles arising from the public good property of information may seriously reduce the incentive to invest in innovations.

While new information is costly to create it is relatively cheap to disseminate or imitate. Thus unless inventors are at least partly protected from imitation their incentive to invest in R&D may be severely curtailed. This problem is exacerbated due to the high degree of risk associated with many R&D projects. Successful projects must generate extra normal profits in order to insure sufficient ex ante prospective profits. However, high profits attract imitators. If imitation costs are low, ex ante profits fall short of the level required for investment. Everyone waits for someone else to invest. The patent system attempts to overcome these difficulties by vesting temporary property rights in the innovator. However, the monopoly so awarded to the innovator raises the price of the information, so that its utilization during the patent period is curtailed below the optimal use. In exchange the public disclosure necessary to obtain a patent ensures that upon expiration the information is available freely. Extension and strengthening of patent rights thus maintains incentives for R&D investment at the cost of postponing some or all of the social benefits from innovation.

Because of this postponement the patent system is, in principle, inferior to alternative mechanisms, such as government financing of all research and development,

making the results available at no cost and raising the revenue via lump sum taxes. Even if lump sum taxes are not available, financing R&D expenditure from general tax revenues (optimally designed) is superior to the patent system. This is because the patent system operates like a benefit tax. Industries are constrained to pay for the research generated in them. While this may be desirable on income distribution grounds, it is not pareto optimal.

The patent system also distorts the allocation of R&D funds among industries and projects. Because it imposes a uniform patent period across all industries. This is not optimal.

However, government finance of R&D in the form of contracts, subsidies or tax concessions suffers from serious informational problems. Governments ability to select agents, direct and monitor their performance is very doubtful. Indeed, there is ample evidence that the productivity of government financed research is considerably lower than that of business financed research. The patent system operates automatically without undue government discretion. As well the availability of appropriation mechanisms such as secrecy, lead time, etc., is very unevenly distributed among industries and products. The patent system equalizes, to some extent, this divergence and therefore supplies a more uniform incentive scheme than other appropriability mechanisms. These are possibly the reasons for its long run survival.

Specials Features of the Pharmaceutical Industry

There are several features of the pharmaceutical industry which are particularly important for the evaluation of patent policy.

A. Patents are of much greater importance as an incentive for R&D in the pharmaceutical industry relative to other industries. In particular, patents

are important for product innovation but much less so for process innovation. Thus the fact that Canadian patent protection is restricted to product by process only,weakens this incentive effect to some extent. As well it tends to distort the direction of R&D effort towards process innovations and towards chemically new products. Thus it does not supply incentives for research into therapeutic properties of relatively cheap natural products.

B. The incentive effect of patent protection in the industry is closely related to the advantage of lead time coupled with high selling expenses to generate brand loyalty. Such selling expenses also increase the speed of diffusion of new drugs,though it may also contribute to their over-use in inappropriate circumstances.

C. The nature of demand in the industry is different than in most industries in two respects. First, demand for many drugs may be highly inelastic allowing very high mark-ups over cost. The monopoly awarded by patent rights is therefore more valuable to the innovator and may lead to lower social benefits relative to the innovator's profits under certain circumstances. Patent policy must take account of this possibility. Second, public insurance covers a significant proportion of consumer drug expenditure. In the absence of appropriate controls this creates a serious moral hazard, making potential demand even more inelastic than it would otherwise be. The evaluation of patent policy must take account of provincial policies designed to remove this moral hazard problem.

Patent protection is currently seriously restricted by the availability of compulsory licensing for both the production and importation of pharmaceuticals. It is important to evaluate the interaction of compulsory licensing,patent protection and the royalty rate.

E. Current government regulation both in Canada and abroad requires extensive testing prior to the use of new pharmaceuticals. This raises significantly the cost of innovation and imitation and causes significant delay in the introduction of new drugs. This reduces potential profitability of R&D in the industry and creates barriers to entry for new innovators, thus reducing potential competition. To the extent that such barriers significantly reduce potential competition, so as to endow some firms with a monopoly in innovation, this significantly affects patent policy. However, on a world-wide basis it is doubtful that a significant monopoly in innovation exists.

F. The pharmaceutical industry competes on a global basis. Most production takes place in very few facilities, generally outside Canada. Because of economies of scale and the small size of the Canadian market, it is unlikely that significant production can take place in Canada.

Optimal Patent Design, a Global Perspective

While Canada's nationalist interest as an individual country may dictate policies which may diverge from those of other countries, its interest as a member of the world community must not be ignored. Canada is a signatory to the Paris convention and the optimal design of globally efficient patent policies is of interest. As well, the determination of the benefits and costs of pursuing an independent patent policy depend critically on the optimality of patent policy elsewhere. In conformity with theoretical literature we evaluate patent policy in terms of the optimal patent life. A critical consideration appears to be the structure of the innovation industry, that is, whether there is a monopoly or competition in innovation. In general, the net social benefits of an innovation

are equal to the present value of (a) gross profits of the inventor, plus (b) the increase in consumer surplus due to potential price reductions or new product availability, less (c) the resource costs of both the successful and unsuccessful innovators, less (d) the administrative costs of the system. (We shall ignore the last). The difference between competition and monopoly arises because in the monopoly case rents to successful innovations are not fully offset by expenses on unsuccessful ones. Excess profits to monopolies is thus a part of the social welfare generated by the innovation. However, in competition there is no excess profit. Where entry into the innovating industry is free, competition for the excess profits of the successful innovator dissipates all these rents in the form of expenses on R&D by unsuccessful innovators. While the increase in R&D effort by unsuccessful innovators may increase the probability of innovation or cause it to advance in time, the marginal productivity of additional entrants is likely to drop very rapidly.

In most industries including the pharmaceutical industry competition is much more likely to prevail than monopoly. Competition in innovation frequently arises not only within an industry but also outside it. We therefore concentrate on the competitive case.

In order to determine the optimal patent life, we must weigh the social losses due to postponement of the consumer surplus in the post patent period against the increased flow of innovations due to the increased profitability from a longer patent life. At the margin, these should just offset each other. The main determinant of the optimal patent life is therefore the rate at which the marginal productivity of additional resources attracted to R&D falls.

An intensive investigation of alternative models describing the R&D process identifies three main reasons for the decline in the marginal productivity of R&D.

First, where probabilities of innovation and its potential profitability are high additional entrants attracted by the prospect of higher profits due to longer patent life are likely to replicate the R&D process of rivals. Thus their marginal productivity is likely to be very low.

Where probabilities of early innovation are low, firms are less likely to replicate rivals research. However, as additional firms enter in response to improved profitability prospects, the likelihood that at least some firms would replicate experiments of some other firms increases. This can be characterized as statistical duplication and its extent rises with the number of firms and the probability of early discovery. As the number of firms entering is likely to be a function of the inherent profitability of the project, marginal productivities are likely to be low and rapidly declining for more profitable projects. For many such projects the patent life is inversely related to profitability. A series of tables, describing optimal patent lives for different processes, reveals that optimal patent life falls rapidly as the inherent profitability of projects increases and as probabilities of discovery rise. For a wide range of projects optimal patent life for process innovations ranges between three months and five years. Only for relatively low profitability, low probability innovations is a patent life of more than five years justified.

An additional element which tends to reduce the productivity of investment in general and of the patent system in particular, occurs due to the "first at the patent office wins" problem. Competition forces acceleration of the R&D process which raises research costs. As well, premature innovation occurs in order to foreclose rivals entry.

The results are somewhat more ambiguous for product innovations. Because consumers benefit from product innovations even during the patent period, the

marginal costs of postponement of some benefits due to increased patent life is not as high and the benefits from additional R&D are higher than in corresponding process innovations. Therefore longer patent lives are optimal. However, the results are sensitive to the specification of demands. For linear demand curves optimal patent periods rise by about 10% for short patent life of about 1 to 2 years and up to about 30% for patent lives over 5 years. However, for constant elasticity demand curves the results depend on the elasticity of demand. As the elasticity of demand falls, patent life rises significantly. Thus, for example, for projects which give an optimal patent life of 1 year for a small process innovation, optimal life almost doubles for a product innovation if demand elasticity is 1.2.. However, optimal patent life only increases by 20% if demand elasticity were 3. These results are even more extreme for longer patent lives. For example, an innovation which yields an optimal patent life of 5 years for a process innovation leads to an infinite patent life if the elasticity of demand is 1.2 and to about 9 years if the elasticity of demand were 3. Because elasticities of demand in the pharmaceutical industry are likely to be low at the monopoly price, as evidenced by the high price cost margins, it is difficult to evaluate these elasticities at very high prices. It is therefore likely that optimal patent life for product innovations is significantly longer than that indicated for process innovations.

However, even in this case it appears that, except for basic research, patent lives should be considerably shorter than the current 17 year period, possibly less than half.

So far our discussion was restricted to patent lives which were specific to the nature of the project in question, however, in general patent lives must be uniformly determined across many projects. This introduces an additional concavity into the process. As patent lives are increased social benefits from profitable and socially

beneficial projects are postponed, but new resources mostly add marginal, relatively unprofitable projects. Thus, the average patent period must be more heavily weighted towards the lower end of the scale, yielding an average optimal patent life which is probably shorter than 5 years, even for product innovations.

The introduction of compusory licensing changes the picture, we show that compulsory licenses with an appropriate royalty rate is always superior to a shortened patent life which yields the same revenues to the innovators. Thus, by retaining the same incentive to R&D the welfare costs due to under-utilization of innovations are reduced. Optimal policy therefore requires very long patent lives coupled with compulsory licensing and appropriate royalties. An investigation of the appropriate royalty rates suggests that they vary between 10 and 30% of the maximum innovator profit for most R&D projects. Again, an average for the industry as a whole is likely to be closer to the lower than to the upper end.

Optimal Patent Life--National Considerations

The preceding analysis utilized the social welfare function which weighted all individuals equally, regardless of their position as producers, inventors or consumers and regardless of their nationality or residence. We may however, wish to evaluate patent policy from a Canadian perspective in terms of the costs and benefits to Canada rather than world-wide. The work of Berkowitz and Kotowitz (1982) suggests that because Canadians are mainly consumers of pharmaceutical products and because the market size is so small relative to world markets, the incentive effects of the Canadian patent system in terms of encouraging additional innovation are extremely small but the costs are high. As very little local production and/or commercial innovation takes place in Canada, the patent system creates a significant transfer from Canadian consumers to foreign multinationals.

Even if such transfers were fully transformed into productive R&D expenditure, the benefits to Canadians from these innovations would be very small, approximately equal to our world consumption share (2%).

If the patent period in the rest of the world were optimal or shorter than optimal, it could be argued that the additional incentive to R&D created by Canadian patents, more or less offset the benefits which we get from the incentives supplied by consumers in other countries. However, if the patent system in the rest of the world is much too long, as is likely the case, increased payments by Canadians for foreign inventions bring forth investment in R&D which mostly duplicates other efforts and the productivity of which is likely to be very low. Thus, while high prices paid to foreign pharmaceutical companies increase the profits of these successful innovators, this increasing profits is largely dissipated by additional unproductive R&D conducted elsewhere in the hope of capturing some of these rents. Thus, these transfers from Canadian consumers accrue mainly to foreign researchers but are unlikely to generate significant social benefits to anyone.

Thus, if such transfers were eliminated or drastically reduced, they would not impose a serious cost on the rest of the world but benefit Canada significantly. When these considerations are coupled with the superiority of a compulsory licensing approach it appears to justify the present system of compulsory licensing with relatively low statutory royalties.

There is however, one possibly important, caveat to the preceeding argument. While revenue from sales or royalties in Canada is not likely to yield significant benefits from additional global research, it does act as a significant incentive for the significant expenditure necessary to allow usage in Canada and for speedy dissemination of information about new drugs in Canada. Thus, some incentive must be given to originators of drugs to disseminate the information

about them quickly and efficiently. A licensing scheme may not do this, because the profit margins may be too low. Therefore, a system of a short patent life, perhaps 2 or 3 years without licensing, followed by a long patent life with compulsory licensing may indeed be superior.

INTRODUCTION

General Rationale For the Patent System

In order to undertake research and development activity firms must have prospects of profits. It is well established that the volume of R&D is positively related to potential profitability. However, there are serious obstacles to the ability of innovators to appropriate to themselves the benefits due to their innovations. These obstacles may seriously reduce the incentive to invest in innovations.

These obstacles arise from the public good property of information. While new information is costly to create it is relatively cheap to deseminate or imitate. Thus, unless inventors are at least partly protected from imitation their incentive to invest in R&D may be severely curtailed. This problem is exacerbated due to the high degree of risk associated with many R&D projects. Successful projects must generate extra-normal profits in order to insure sufficient prospective profits. Thus, the prospect of high profits from successful projects must be sufficient to compensate investors for the unsuccessful ones. However, high profit attracts imitators which reduce profits from successful projects. If imitation costs are low, the ex ante profits prospects of R&D fall short of the level required for investment. Everyone waits for someone else to invest (Spence, 1984).

The patent system attempts to overcome these difficulties by vesting temporary property rights for the invention in the innovator enabling him to use it exclusively or sell it to others. However, the monopoly so awarded to the innovator raises the price of the information so that its utilization

during the patent period is curtailed below the optimal use. In exchange the public disclosure due to the patent insures that upon expiration the information can be optimally used. Moreover, during the patent period monopoly pricing may still leave consumers some benefits in the case of a large process or product innovation.

However, the patent system is not the only mechanism of appropriation available to inventors. Indeed many studies suggest that other mechanisms may be more important. In particular, in many industries secrecy is a viable and preferred alternative, and in other cases lead time coupled with complementary investments in marketing may be sufficient. Although patents may enhance the efficiency of such other mechanisms. A recent large scale survey of U.S. corporations (Levine et al., 1984) found that patents were considered of least importance for process innovations relative to lead time,superior sales, and service effort, moving quickly down the learning curve and secrecy. While patents were somewhat more important for new product innovations they were still less important than most other means of appropriation except secrecy. However, the survey revealed that patents are particularly important in protecting new products in the pharmaceutical and chemical industries. Patents are therefore mainly a compliment to a range of instruments available for appropriation. Although they may also serve as a substitute for sales and service and other costly instruments. However, patents are clearly intended as a substitute for secrecy because of the disclosure requirements to obtain a patent. Indeed, this is one of the main advantages claimed for patents.

The significance of the advantages due to disclosure depend on two related questions: (a) Are patents an effective substitute for secrecy; and (b) does patent disclosure help significantly in the dissemination of the information.

The studies by Levine et al. and others suggest that where secrecy is an effective means of appropriation patents are not effective. Thus, in process innovations reliance on secrecy and complementary instruments appears to dominate patents. However, in product innovation where secrecy is much more difficult to maintain patents are considerably more important. The reason for the superiority of secrecy in the case of process innovations are difficulties in obtaining patents for many processes and ease of circumventing the patent once the principle is known. Note however, that where patent protection is effective patents are of greater value to small innovators than to large ones because they enable them to license their innovations for use by others. A course of action which can not be undertaken when secrecy must be relied on.

The patent system may help in the utilization of the innovation and the dissemination of information both during and after the patent expiry. The disclosure of an invention through the patent grant may give ideas to innovators in other industries through publication in the official gazette. It may thus foster additional technical advance. Second the monopoly right accorded to the innovator by the patent yields a strong incentive to disseminate information about the innovation in order to maximize the demand for it. To the extent that the innovator is protected from imitators, it is advantageous for him to disseminate information about the innovation without fear of conferring external benefits on competitors. Lefler (1981) and Telser (1975) document the existence of such instances for the drug industry. A second avenue of the dissemination of technical information occurs through licensing. Levine et al., document the importance of licensing in the transmission of technological information from innovators to users and to other researchers. The combination of disclosure

through the patent and information dissemination through advertising and licensing, increase the probability of utilization of innovations by competitors after the patents expiry. However, to the extent that patents enable corporations to create significant barriers to entry through successive patenting, advertising and brand loyalty, etc., post patent utilization of the innovation may not increase.

The degree to which the diffusion of new technology during and in the post patent period is enhanced by patents remains an open question. Many critics argue that the information provided by patents is inadequate to commercially exploit post inventions. The Cambridge report for example, asserted that patent disclosures do not provide all the knowledge necessary to exploit an invention commercially and thus falls short as an incentive to disclose. In addition, Shearer (1970) argued that "even when patent protection is sought in exchange for disclosure the details of an invention may be with luck kept secret until the patent is issued, in the United States 3 years on the average after an application is filed.

It is clear from our discussion that patents must be evaluated within the complete system of appropriation of innovation benefits. The patent system acts mainly as a compliment but also as a substitute for some of these other mechanisms. Because many of these other mechanisms such as sales and service, efforts, have real resource costs any reductions in these instruments due to the existence of patents lead to resource savings and therefore to possible social benefits. However, such improvements due to the patent system are uncertain because of the common pool problem.

The common pool problem arises from competition in innovation. Innovative innovators in competing for the rents accruing to the successful

innovator tend to dissipate such rents by excessive investment in R&D. This occurs because resources are allocated to R&D until the average product of R&D investment rather than marginal product equals marginal resource costs. The greater is the divergence between average and marginal returns the greater the resource dissipation. The patent system by creating large rents to the successful innovator leads many competitors to invest in the hope of capturing the rent. If there are no barriers to entry into the innovation industry expected profits in any line of endeavour are expected to remain normal. Thus during the patent period there may be no social benefits from the innovation. The results are similar for other appropriation methods. To the extent that they guarantee higher profits to successful innovators they generate excessive investment to capture such profits. The degree of such potential over-investment and possible waste is directly dependent on the concavity of the overall innovation process. That is, on the difference between the average and the marginal product of R&D. As firms invest in R&D until the average product equals marginal resource costs the excess of average over marginal product leads firms to invest in projects the marginal product of which is smaller than the marginal resource costs. It is this consideration which makes the patent system as well as other rent-generating mechanisms socially inefficient in generating optimal research effort. It must, however, be stressed that this problem is common to all rent generating appropriation mechanism and not only to patents. The patent system must therefore be judged in the context of the existence of alternative instruments such as secrecy, lead time, sales efforts, etc. to generate appropriable rents. Unless alternative mechanisms to encourage innovation remove these other means of appropriation their effect will be very uneven. If they are applied across the board such as in the case of subsidies

or taxes, they will reinforce the tendencies for over-investment where such instruments are useful without necessarily inducing the appropriate investment in those industries where other appropriation mechanisms are weak. As alternatives which have been suggested to the patent system such as government contracts and subsidies to research and development must be applied in a discretionary manner.

If it were possible to eliminate all expropriation mechanisms by the government financing all research and development while making the results available at no cost while raising the revenue via lump sum taxes the result would be pareto optimal. However, these are not viable options. Lump sum taxes are not available so that raising the revenues required to finance such a scheme would cause distortions which may be as great as those eliminated from the R&D areas. As well government contracts or subsidies require the government to direct research activity a role it is singularly unfit to do. The alternative course of selective contracts or subsidies available only to industries or projects where other mechanisms of appropriability are very weak, suffers from the same problem. The degree of knowledge which such discretion requires is most unlikely to be possessed by the government. Indeed, there is ample evidence that the productivity of government financed research is considerably lower than that which is financed by business. It thus appears that instruments which improve appropriation make it more uniformly available across industries and projects must be the mainstay of government policy.

An instrument which we have not considered so far is compulsory licensing. Obviously such an instrument is only useful where protection is effective in preventing imitation. Thus compulsory licensing must be considered together with the general patent policy. The advantage of compulsory licensing

is that a longer patent life with compulsory licensing at an appropriate price is always superior to a shorter patent life which yields the same revenues to the innovator. Thus, by retaining the same incentive to R&D the welfare costs due to under-utilization of innovations are reduced. Of course, this raises the issue of the appropriate licensing fee. However, in principle this is not different from the determination of optimal patent life. Thus, it is clear that the issue of compulsory licensing should be explored further in the context of the design of optimal patent systems. This is particularly important in the context of the pharmaceutical industry in Canada where compulsory licensing has come under severe attack.

Special Features of the Pharmaceutical Industry

There are several features of the pharmaceutical industry which are particularly important for the evaluation of patent policy.

A. Patents are of much greater importance as an incentive for R&D in the pharmaceutical industry relative to other industries. In particular, patents are important for product innovation but much less so for process innovation. Thus the fact that Canadian patent protection is restricted to product by process only, weakens this incentive effect to some extent. As well it tends to distort the direction of R&D effort towards process innovations and towards chemically new products. Thus it does not supply incentives for research into therapeutic properties of relatively cheap natural products.

B. The incentive effort of patent protection in the industry is closely related to the advantage of lead time couples with high selling expenses to generate brand loyalty. Such selling expenses also increase the speed of diffusion of new drugs, though it may also contribute to their over-use in inappropriate circumstances.

C. The nature of demand in the industry is different than in most industries in two respects. First, demand for many drugs may be highly inelastic allowing very high mark-ups over cost. The monopoly awarded by patent rights is therefore more valuable to the innovator and may lead to lower social benefits relative to the innovator's profits under certain circumstances. Patent policy must take account of this possibility. Second, public insurance covers a significant proportion of consumer drug expenditure. In the absence of appropriate controls this creates a serious moral hazard, making potential demand even more inelastic than it would otherwise be. The evaluation of patent policy must take account of provincial policies designed to remove this moral hazard problem.

Patent protection is currently seriously restricted by the availability of compulsory licensing for both the production and importation of pharmaceuticals. It is important to evaluate the interaction of compulsory licensing, patent protection and the royalty rate.

E. Current government regulation both in Canada and abroad requires extensive testing prior to the use of new pharmaceuticals. This raises significantly the cost of innovation and imitation and causes significant delay in the introduction of new drugs. This reduces potential profitability of R&D in the industry and creates barriers to entry for new innovators, thus reducing potential competition. To the extent that such barriers significantly reduce potential competition, so as to endow some firms with a monopoly in innovation, this significantly affects patent policy. However, on a world-wide basis it is doubtful that a significant monopoly in innovation exists.

F. The pharmaceutical industry competes on a global basis. Most production takes place in very few facilities, generally outside Canada. Because of economies of scale and the small size of the Canadian market, it is unlikely that significant production can take place in Canada.

OPTIMAL PATENTS FROM GLOBAL PERSPECTIVES

Basic Criteria

Following Stiglitz (1969) we shall attempt to evaluate the special features of the patent system with respect to: (a) that total economy-wide allocation of resources to research and development; (b) the allocation of R&D among alternative industries and projects; (c) the total dead-weight loss resulting from the divergence between price and marginal cost.

In the context of these questions we shall attempt to evaluate the factors which may be important for the improvement in the performance of the patent system. In order to simplify the analysis and focus on the key elements we shall abstract from complications arising from the existence of other appropriation mechanisms and from the shortcomings of the patent system in protecting property rights. We shall thus assume that patents are uniformly effective in protecting property rights during the patent period enabling us to concentrate on the issue of patent life. As well, we shall investigate the optimal combination of life and royalty fees in the presence of compulsory licensing. In this chapter we adopt a global international view for evaluating the patent system. While Canada's nationalist interest as an individual country may dictate policies which may diverge from those of other countries its interest as a member of the world community must not be ignored. First, Canada is a signatory to the Paris Convention and the optimal design of globally efficient patent policies is of interest. Second, the determination of the benefits and costs of pursuing an independent patent or licensing policy depend critically on the optimality of patent policy elsewhere.

In order to analyze the costs and benefits of the patent system it is first necessary to evaluate the net benefits derived from its product: invention and innovation. The net benefits of an invention are generally measured in terms of a social welfare function, which is generally expressed in terms of consumer and producer surplus. Thus valuation is based on the market valuation of the product produced by the invention, aggregated over all consumers and producers without regard to income distribution. The net social benefits of an innovation are equal to the present value of (a) gross profits of the inventor, plus (b) the increase in comsumers surplus due to potential price reductions or new product availability, plus (c) the change in producers surplus due to increased demand for the services of some specialized resources, less (d) the resource costs of both the successful and the unsuccessful innovators, less (e) the administrative costs of the system. In addition, adjustments should be made for externalities, increasing the net social benefit for non-internalized economies and reducing net social benefit by the diseconomies overlooked by the market. Generating the common pool problem mentioned before. As we show later, the nature of competition in the innovation industry leads to drastic differences in the optimal patent policies. We believe that competitive free-entry innovation industry more closely corresponds to reality than the monopoly or even the monopoly or even oligopoly model. Competition in innovation frequently arises not only within an industry but also outside it. As well competition is global and there are very few industries with significant monopoly power world-wide. This is clearly the case with the pharmaceutical industry which comprises a large number of firms competing for innovations on a global basis. We shall therefore concentrate our analysis on the competitive case.

Following the theoretical literature (Nordhaus, 1969, 1978) we assume that factor supplies are infinite and ignores the administrative costs and externalities of the patent system. The main focus is on the benefits to the inventor and consumers. We therefore distinguish between three types of invention.

(a) A minor process invention in a competitive industry.

(b) A major process invention in a competitive industry.

(c) Product inventions.

Each of these leads to a somewhat idfferent confirguration of net benefits. We shall analyze each in turn.

A. A Minor Process Invention in a Competitive Industry

This is an innovation which reduces costs only slightly, in an industry which is competitive prior to the innovation. The situation is illustrated in Figure 1 where C_0 are the costs prior to the innovation, and C_1 are the costs after the innovation. The maximum royalty which the innovator can charge for licensing all producers in the industry is equal to the total cost saving due to the innovation, at the pre-innovation level of output. This is because the innovator faces competition, or at least potential competition, from unlicensed firms using the old technology and, hence, can not raise the price above that which prevailed prior to the innovation.

Thus the benefits to the innovator are equivalent to the full cost saving due to the innovation, for the period over which he can exercise a monopoly on this innovation. During that period, consumers do not benefit from this innovation, as there is no decrease in the price of the product. Other factors of production, however, may benefit or suffer from the innovation,

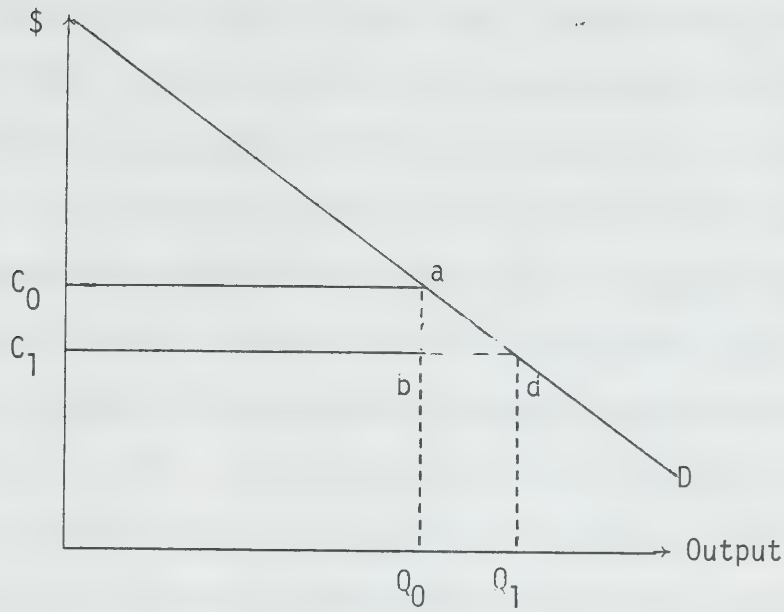


Figure 1

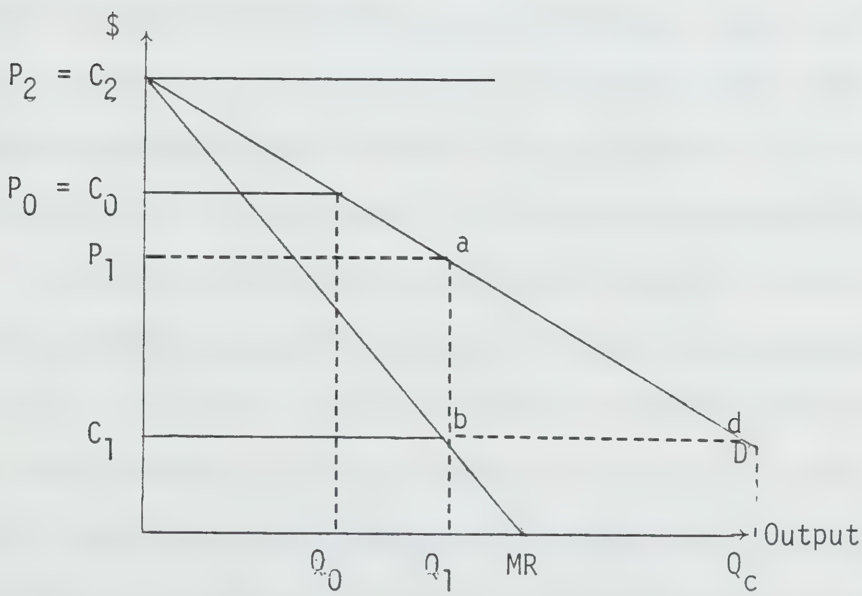


Figure 2

to the extent that it changed relative demand and supply for factors of production and that the supply of these was not perfectly elastic.

In the period after the expiry of the monopoly over the innovation, competition will drive the price of the product down to C_1, eliminating the monopoly profits of the innovator. The reduction in price will, therefore, increase consumer surplus by the full cost savings due to the innovation, which are equal to the monopoly profit, plus the area abd which represents the gain in consumer surplus from the increased output at the lower price.

It is this consumer surplus, representing the excess of social benefit over the private benefit to the innovator, which is at the heart of the dilemma of the patent system. An increase in the scope and period of protection to the innovator increases the private benefits to the innovator and hence, his incentive to innovate, but decreases, or rather postpones, the excess of social over private benefits. Thus optimal patent protection requires that such protection stop short of a perpetual pattern. At the margin, the incentive effects of additional patent protection must be balanced against the loss in consumer surplus which arises from it.

To arrive at the net profit of the patent system, we must deduct from the present value of these gross benefits to the innovator and consumers the resource costs incurred in the invention and dissemination process.

In order to analyse the net benefits of the patent system, it is necessary to examine the industrial structure in the innovating industry as well. Nordhaus (1969, 1972) assumed a monopoly in the innovation industry, i.e. a single innovator. Under these circumstances, the innovator will invest resources in research and development up to the point, where the expected present value of the marginal royalty from from an increment to investment is

just equal to the marginal cost of that investment. If, as is generally assumed, the function relating such benefits and costs of innovation is concave, the present value of the royalties will exceed the resource costs. Thus, the net private benefits to the innovator will be positive.

However, if competition prevails in the innovation industry the resource costs of unsuccessful innovators must also be taken into account. Competition in the innovation industry will cause total resource costs expended in innovation to equal the present value of the royalties to innovation.

B. A Large Process Innovation in a Competitive Industry

Such a process innovation is assumed to lower the costs of production sufficiently below those prevailing prior to the innovation, so that the optimal royalty payment falls short of the cost savings due to the innovation. This can be seen in Figure 2, when costs fall from C_0 to C_1, maximum industry profits occur at P_1 which is lower than C_0. In this case, the gain to the inventor falls short of the cost decrease due to the innovation. Some of the gain accrues to consumers who benefit from the lower price even during the periodof the patent monopoly. Thus the excess of social benefits over private benefits in this case is greater than in the small innovation case.

We shall concentrate on the two extreme cases of process innovation and product innovation as they bracket the large process innovation.

C. Product Innovation

Where a new and superior consumer product is introduced, Usher (1964) demonstrated that, even when the product is priced monopolistically, it will confer upon consumers a net consumer surplus. Thus, the benefits to society

from a product innovation are equal to the monopoly profits of the innovator plus the consumer surplus represented by the area under the demand curve for the new product above the monopoly price. Upon expiration of the patent, consumer surplus will increase to encompass the monopoly profit, plus the additional consumer surplus generated by the increased output.

It is useful to think of a product innovation as a large process innovation in which the pre-innovation cost intersects the demand function at zero output. This is illustrated in Figure 2 where the pre-innovation cost is C_2 and the post-innovation cost is C_1. It follows that the conclusions for a large process innovation would apply to the case of product innovation, but in more drastic form.

A detailed analysis of the optimal patent life for individual projects was developed by Nordhaus (1969) for the case of small process innovations under monopoly in innovation. He found that the optimal patent life varies considerably, depending on the parameters of the function relating research and development to cost savings, and on the elasticity of demand for the product. In particular, the optimal patent life is shortened as the elasticity of demand and the profitability invention rises so that patent life should be shorter for industries in which technological conditions favour invention and which have elastic demand. However, departures from optimal patent life do not affect welfare significantly, so that patent life of from 5 to 1,000 years are almost equally efficient. Therefore the social losses due to a uniform patent life are not significant. However, for high elasticity of demand and relatively large or important inventions, the efficiency of the patent system relative to an ideal system of lump sum taxes and subsidies is quite low. Thus the patent system is efficient for small inventions where demand is relatively inelastic but inefficient for larger inventions where demand is elastic, regardless of the patent life (within broad limits).

Two major conclusions emerged from Nordhaus' analysis. First, a fixed patent life is not optimal in theory, although it may be unavoidable in practice. If we are to err on one side, the analysis suggests too long a patent life is better than too short a patent life. For run-of-the-mill innovations, the losses from monopoly are small compared to the gains from invention. The best way to prevent abuse is to ensure that trivial inventions do not receive patents.

Second, the complications arising from risk, drastic inventions, imperfect product markets and 'inventing around' patents, generally point to a longer rather than shorter patent life.

However, these conclusions critically depend on the assumption of monopoly in innovation. This analysis eliminates the common pool problem, and any welfare losses which arise due to competition for the rents. Increased patent periods, which raise rents to successful innovators are therefore part of the welfare gain from innovation.

However, competition in innovation tends to dissipate such rents. Increased rents to successful innovators attracts a greater amount of R&D resources, but the marginal productivity of such additional resources may be rather low because of possible statistical duplication, actual replication and acceleration of research program. This problem is particularly severe for inframarginal projects, which would have been profitable at short patent life. Thus the incentive effects of the patent system, while effective in bringing forth more R&D resources, may be ineffective in bringing forth more innovation.

Because we believe that competition in innovation is almost universal, it is important to investigate the degree to which these problems of duplication, replication and acceleration are likely to arise and the extent to which they

are likely to modify Nordhaus' conclusions. To accomplish this we shall sketch out a few simple models which highlight each of these features in turn and derive the relevant optimal patent periods.

The Nature of the Innovation Process

The nature of the processes assumed will illustrate the general principles and factors affecting optimal patent life for individual projects, for specific industries and the economy as a whole. In particular, we shall attempt to identify the factors which determine the nature and degree of concavity of innovation processes and thus interaction with product and industry characteristics.

Model A

An innovation can be pursued along several potential research avenues (N) only a subset (S) of which lead to innovation. Potential inventors are assumed to have an accurate idea of $S/N = p$ the probability of choosing a successful avenue, but do not necessarily know all the avenues. For simplicity, we assume that potential inventors cannot distinguish the relative potential profitability of avenues and hence treat them as equally profitable.* Investment of x_i in each avenue generates increased profit if the innovation is obtained. Thus the annual profit from a successful innovation is a function of the investment level undertaken by the relevant inventor in the relevant avenue.

(1) $h_i = h(x_i)$, $h' > 0$, $h'' > 0$ if $x < x_0$ and $h'' \leq 0$ if $x \geq x_0$.**

Because competition is known to exist, investment takes place at maximum speed, so as to obtain prior discovery. Further acceleration of R&D cannot be obtained by increased expenditure. As the probability of discovery depends only on the choice of avenue, increased expenditure cannot increase it.

---

*This assumption is not essential, but simplified the analysis.

**$h(x)$ is the area c_0c_1ba in Figure 1.

The first firm to discovery obtains a patent which blocks all other firms. The date of discovery is random within very narrow limits, which do not affect the present value of the innovation but determine priority and hence, who receives the patent. The probability of innovation given n innovators is $b(n) \leq pn \leq 1$; $b' \leq p$, $b'' \leq 0$.

Define (2) $\psi = (1-e^{-rT})$ where T is the patent period, and r is a discount rate. Then the expected present value of the innovation to an innovator when there are n innovators is

(3) $\int_0^T g(n)h(x_i)e^{-rt}-(x_i)= \psi\, g(n)h(x_i)-x_i$ where $g(n) = b(n)/n$ is the probability of prior discovery by inventor i, equal to the probability of discovery $b(n)$ times the probability that inventor i was the first at the patent office given that a discovery occurred $(\frac{1}{n})$.*

The specification above is general enough to allow us to consider a variety of possibilities of duplication in research effort. Full duplication (replication is represented by $b(n) = p = 1$. This will be the case where all research avenues lead to successful innovation, so that firms simply duplicate each others efforts. This is most likely to be the case for development efforts,

---

*We ignore the delay in innovation past the investment, as it does not affect the results.

where technical uncertainty is small--the processes required for successful implementation are fairly well known and probability of technical success high.

In the pharmaceutical industry case, this situation is likely to occur during the clinical stages of the research process to determine dosages, safety, etc.

The same result of complete duplication occurs where only one avenue is successful and firms know rivals' investment avenues, so invest in unexplored avenues, if any are available because all avenues are equally likely, all will be covered. Alternatively each firm invests in all avenues, so that $b(n) = 1$. In either case if profitability is high enough, additional firms enter, duplicating the effort of others.

If firms are unaware of rival choice of avenue, or if the number of avenues is very large, some avenues may be duplicated by entering firms and others left unexplored altogether. The marginal productivity of additional entrants therefore depends on the probability of entering an unexplored avenue and the probability that the avenue chosen is successful. It is clear that the first is a rapidly declining function of the ratio of firms to avenues (n/N), assuming each firm enters one avenue. Similar results occur when a firm may enter more than one avenue, if the number of potentially succesful avenues is large. Thus the cause of such declining marginal productivity is statistical duplication.*

Each firm is assumed to maximize profits given the research strategy of its competitors, i.e. it chooses x_i to maximize expected profits from research, given rivals' number and research outlays. This yields the first order condition, which determines x - the research outlay.

---

*An additional cause for falling marginal productivity may arise if _a priori_ some avenues are more promising than others. In this case firms tend to crowd the promising avenues, leaving the less promising ones to later entrants or not exploring them altogether. This source of concavity is similar to that which results from differential profitability and will be explored later.

(4) $\psi b(n)h' = rn$.

If entry into innovation is free, such entry occurs as long as expected profits to each firm from R&D into a given area are in excess of normal profits. Thus for each firm, industry equilibrium implies

(5) $\psi b(n)h(x)/rn = x$.

combining (4) and (5) yields:

(6) $h' = h/x$.

Equation (6) can be solved for x independent of n and ψ , so that the equilibrium R&D investment by each firm (in each avenue) - x^* is fixed by technology and independent of patent policy. Figure 3 illustrates this solution. At x^*, h'-the slope of $h(x)$ equals the ratio of $h(x)/x$ - the average product of x.

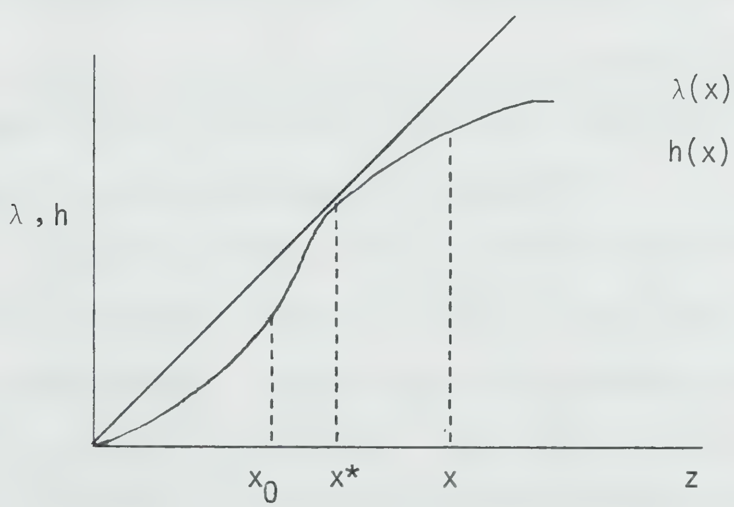


Figure 3.

Note that this is not the scale which would be chosen by a monopolist controlling R&D investment in all avenues, or if the number of competitors were different than that determined by the zero excess profit conditions. Indeed,

each firm will wish to adjust its scale as the patent period rises, but entry of new firms due to increased potential rents causes it to remain at x*. The effects of changes in the patent period (represented by ψ) occur in the competitive case, via changes in the number of avenues explored, leaving investment in each avenue (x*) unchanged. Given x*, the number of avenues explored adjusts to changes in ψ, so as to leave $\psi\, g(n)$ unchanged.

We are now in position to determine optimal patent life for individual projects, characterised in terms of the process generating the function b(n), intrinsic social and private profitability and the nature of the innovation (process or product).

Small Process Innovation

Consider first a small process innovation, so that product price remains at the pre innovation level during patent life and falls by the full cost savings at the expiration of the patent.

Expected social benefits due to the innovation equal the expected present value of excess profits to innovators plus the expected value of consumer's surplus due to price reduction before and after patent expiration. Competition in innovation implies zero expected excess profit to innovators. Thus net social benefits are equal to the present value of consumers' surplus due to price reduction in the post patent period (as illustrated in Figure 1). If we express all benefits and expenditure relative to initial product revenues (P_0Q_0), h(x) becomes the proportional annual cost savings due to innovation as a function of investment as a proportion of initial product revenue, (i.e. the proportional cost savings due to the innovation).

Then expected social benefits can be approximated by the expression:

(7) $$s = b(n)P_0Q_0 \int_T^{\infty} (h+\eta h^2/2)e^{-rT}dt/r = b(n)P_0Q_0(1-\psi)(h+\eta h^2/2)/r \ .$$

where η is the demand elasticity.

Optimal patent life can then be found by maximizing equation (7) subject to equation (5) and the constraint $b(n) \leq 1$. Recall that x is fixed for any project so $h(x)$ is also fixed hence define $D \equiv P_0Q_0(h+yh^2/2)/r$. To solve, maximize the Lagrangian

(8) $$L = (1-\psi)b(n)D+\lambda(1-b(n))$$

with respect to ψ.*

The Kuhn-Tucker conditions for a maximum are

(9) $$\begin{aligned} &-b(n)D+(1-\psi)Db'dn/d\psi \leq \lambda b'dn/d\psi \\ &\lambda(1-b(n)) = 0 \ . \end{aligned}$$

from (5) $$\frac{dn}{d\psi} = \frac{bh}{rx-\psi b'h} = \frac{bh}{rh/h'-\psi b'h} = \frac{bh'}{r-\psi b'h'} = \frac{bh'}{\frac{\psi bh'}{n} - \psi b'h'} = \frac{n}{\psi\left(\frac{1-b'n}{b}\right)}$$

substituting in (8), and solving for ψ yields

(10) $$\psi = b'n/b \qquad \text{if } b(n) < 1 \ , \text{ i.e. } \lambda = 0 \ .$$

(10a) $$\psi = \bar{n}\, rx^*/h(x^*) \leq 1 \qquad \text{where } \bar{n} \text{ solves } b(\bar{n}) = p\bar{n} = 1.$$

The first term of equation (8) defines the social loss due to postponement of consumer surplus arising from extension of patent life. The second term identifies the increased expectation of social benefits $(1-\psi)D$ arising from the increased probability of discovery (b') due to additional firms investing in R&D in response to increased patent life ($dn/d\psi$). Where the probability of

---

*It is convenient to work with $\psi=1-e^{-rT}$, rather than with patent life T. Note that $d\psi/dT>0$ and $\psi=0$ for T=0, $\psi=1$ for $T = \infty$.

discovery improves in response to entry $(b'>0,\ b<1)$, $\lambda=0$ so patent life is determined so as to equate these marginal costs and benefits, yielding equation (9).

To interpret equation (10), define $y = b(1)h(x^*)/rx^*$, = the expected present value of the average gross expected return per investment if patent life were infinite $(\psi=1)$ and no more than one firm invests in any avenue.* Thus y measures the basic intent profitability of investment in the project condition (10a) may then be rewritten as:

(10b) $\psi = 1/y$.

i.e. patent life is inversely related to basic profitability. When $b(n) = pn$, replication will occur if patent life long and profitability high. ψ is set so that only one firm would enter each avenue avoiding the social losses from unnecessary postponement of patent life. Thus ψ must be set at 1 or at the level which will guarantee normal profit, where $\frac{p\psi h(x^*)}{r} = x$, that is where the average expected industry return is just enough to yield normal profit when no duplication occurs.

In order to solve equation (9) or (10) we must specify the nature of the process determining $b(n)$. Assume, following Tandon (1983), that there exist N avenues to reach the innovation, of which s are successful. Thus the probability of discovery is $b(n) = 1 - (\frac{N-S}{N})^n \equiv 1 - q^n$, i.e. one minus the probability that no one discovered, where research avenues are takne at random without knowledge of rivals' choice of avenues.

Substituting in equation (9) yields:

(11) $$\psi = \frac{(-n\ \ln q q^n)}{1-q^n}$$

---

*Note that $y/b(1)$ is the maximum number of single avenue firms which could profitably pursue the project <u>if</u> each knew the avenues pursued by rivals. Alternatively y is the maximum number of firms, each purusing all avenues.

and

(12) $T = (-\ln[1+n\ln q q^n/(1-q^n)])/r$.

However, n - the number of firms (pursuing one avenue each) engaged in R&D on the project depends on patent life, the probability structure and the inherent profitability. To solve for n we may substitute in $y = (1-q)h/rx$ equation (5) and solve for ψ. Together with equation (11) this yields

(13) $$\psi = \frac{(1-q)n}{y(1-q^n)} = \frac{-n\ln q q^n}{1-q^n} .$$

Solving for n, given optimal patent life yields

(14) $n = \ln[(1-q^{\psi})/-y\ln q]/\ln q$.

From which, by substitution into (12) or (13) we may obtain optimal ψ and T. Table 1 illustrates the combinations of ψ and patent life (T) and the resulting number of firms for combinations of the probability of discovery (1-q) and profitability (y), assuming $r = .2$. The patent period falls as the degree of uncertainty falls and as the number of competitors engaged in a given line of R&D rises. The number of companies is directly dependent on the expected profitability of any research project, that is, on the expected cost saving per dollar of investment in R&D. More productive projects attract a larger number of competitors which leads to at least some duplication and social waste. Hence, more productive projects require a shorter patent life. As well, given expected profitability of each avenue (1-q), the number of firms pursuing R&D towards any specific innovation is likely to depend positively on the number of avenues (N). Thus, the more ways there are of discovering the innovation the greater the number of firms operating and hence, the shorter the optimal patent period.

Table 1

Patent life and number of firms.

Competition with fixed scale

small process innovation

Model A

y / 1-q	2			3			5			10		
	T	ψ	n	T	ψ	n	T	ψ	n	T	ψ	n
.2	5.2	.65	3.6	3.5	.5	5.4	2.4	.37	7.8	1.4	.23	10.8
.5	4.2	.57	1.5	3.0	.45	2.1	1.9	.32	2.8	1.1	.20	3.8
.7	3.6	.51	1	2.5	.39	1.4	1.6	.28	1.8	1.0	.18	2.4

It is clear from Table 1 that, except for projects with very low profitability and low probabilities of success in any of only a few possible avenues, a short patent life is desirable. Such projects are likely to be ones involving very basic research, in which the probability of success is very low because only one of many possible avenues may lead to success and potential cost savings are small.

An alternative may to gain insight into the relevant numbers is to define m as the maximum number of firms which could pursue R&D profitably in a given project, if patent life were infinite $(\psi = 1)$. From equation (5) we have:

$$(15) \qquad m/(1-q^m) = h(x)/rx = y/(1-q)$$

equation (13) may then be written as

$$(16) \qquad \psi = \frac{(1-q^m)n}{(1-q^n)m} = \frac{-n\ln q q^n}{1-q^n}$$

from which we may solve for n, ψ and T as a function of $(1-q)$ and m. As current patent life $T = 17$ implies $\psi = .967$ (assuming $r = .2$), m approximates the number of avenues explored for any project under current conditions. Thus m represents the maximum number of firms which would operate in the industry if each pursued one avenue only. Where firms pursue several avenues the number of firms pursuing a given project may be considerably smaller,* obviously m is closely related to profitability (y). The results are exhibited in Table 2.

---

*As well the probability structure is somewhat different.

Table 2

Patent life and number of firms.

Competition with fixed scale

small process innovation

1-q \ m	3		5		10	
	T	n	T	n	T	n
.2	9.2	1.33	7	2.2	5	3.6
.5	5.3	1.15	3.6	1.8	1.9	2.8
.7	3.0	1.08	2.07	1.5	1.3	2.5

The results confirm the findings of Table 1: unless potential profitability is very low, optimal patent life is much shorter than the current 17 year period. An important corollary is that the optimal number of firms operating under the optimal patent period is relatively small even where potential profitability is very high. Thus our assumptions of free entry appear to be valid. There is no reason to believe that the number of potential R&D competitorsis smaller than the optimal number given optimal patent life, although it may well fall short of the maximum which can profitably operate in any R&D line (m) under the current 17 year period.

Product Innovation

The calculus of private and social returns is somewhat more complex for the case of product innovation. Assuming straight line demand curves as in Figure 2, we define

(17) $$h(x) = \frac{P_2 - C_1}{P_2}$$

is the relevant proportional cost savings due to the innovation. The post patent competitive output is thus

(18) $$Q_c = -\gamma h(x) = (\partial Q/\partial P)h(x)$$

the inverse slope of the demand curve. Output during the patent period

(19) $$Q_1 = Q_c/2 = \gamma h P_2/2 \quad *$$

hence

(20) $$P_1 - C_1 = P_2 - P_2 = hP_2/2 \ ,$$

*MR curve bisects the horizontal distance between the demand curve and output.

where P_1 is the price during the patent period.

Hence innovators annual profits during the patent period

(21) $\pi = (P_1 - C_1)Q_1 = P_2 \gamma h^2/4$.

The zero profit condition for theproject is

(22) $P_2\ b(n)\psi\gamma h^2/4r = xn$, and defining $P_2\ \gamma h^2/4r \equiv G\ /b(1)$ yields

(22a) $b(n)\psi G = xnb(1)$

Social welfare is composed of the discounted value of consumers surplus during the patent period $(P^1 - P_1)Q_1/2$ and the post patent period $(P^1 - C_1)Q_c/2$.

(23) $S = P_2 b(n)[\psi\gamma h^2/8 + (1-\psi)\gamma h^2/2]/2 = P_2\ b(n)\gamma h^2(4-3\psi)/8r \equiv b(n)G(4-3\psi)/2$.

Maximizing S subject to equation (22a) yields (assuming b < 1)

(24) $-3b(n)G/2 + b'G(4-3\psi)\frac{dn}{d\psi}/2 = 0$

and substitutiong from (22) for

(25) $\frac{dn}{d\psi} = \frac{nb}{\psi(b - b'n)}$

yields

(26) $\psi_p = \frac{4b'n}{3b}$ where the subscript p indicates product ( ).

Following our previous assumption about profitability, we may define profitability as $y = b(1)G/x$; then from (22a)

(27) $p = \frac{n(1-q)}{y(1-q^n)} = \frac{4(-n\ln qq^n)}{3(1-q^n)}$

solving for n yields

(28) $n_p = \{\ln.75+\ln[(1-q)/-y\ln q]\}/\ln q = \ln.75/\ln q+\ln[(1-q)/-y\ln q]/\ln q = \ln.75/\ln q+n>n$

and substituting for n in equation (27) we obtain optimal ψ_p. It can be immediately seen from equation (26) that $\psi_p < 4\psi/3$ as b is concave in n. Thus the maximum increase in ψ is about 30%. However, for most cases the increase is very much smaller. Table 3 illustrates optimal ψ_p, n_p and the corresponding patent life for linear demand. A comparison of Tables 3 and 1 reveals that in general, ψ_p falls short of 1.15ψ and similarly $n_p < 1.15n$. The corresponding patent lives vary in a similar manner. However, note that for low profitability-low probability projects, which require $\psi > .5$, a proportional increase in ψ leads to a greater proportional increase in patent life. Thus for marginal projects, optimal patent life for product innovations may be considerably longer than that of equivalent profitability process innovations. This is because, during the patent, consumers benefit from product innovation even if very marginal, but do not benefit from a small process innovation. Therefore, marginal process innovations which require a very long patent to make them profitable even to one innovater, confer almost no benefits on society at large, as the present value of the small benefits the project generates, is drastically reduced by postponement. However, even for marginal projects with long patent life, for product innovation, some benefits are available to consumers during the patent, justifying a longer patent life.*

---

*Note also that these results are restricted to the analysis of linear demand curves. As we shall show later in analysing model B, patent life is significantly longer if demand curves are log linear.

Table 3

Patent life competition with fixed scale (r=.2)

product innovation

1-q \ y	2			3			5			10		
	T	ψ	n	T	ψ	n	T	ψ	n	T	ψ	n
.2	6.7	.74	4.9	4.2	.57	6.6	2.7	.42	9.1	1.5	.26	12.1
.5	5.2	.65	1.9	3.4	.49	2.4	2.3	.36	3.2	1.4	.22	4.2
.7	4.0	.55	1.9	2.9	.44	1.6	1.9	.31	2.0	1.0	.19	2.6

For most projects, optimal product patent lives are only slightly longer than those for small process innovations. Even for low probability-low profitability events they fall considerably short of the current patent period.

These results are even stronger for projects in which actual replication is likely. From equation (10a) we see that optimal ψ falls short of .5 for all projects where $y \geq 2$. Implying patent life shorter than 3.5 years, even for relatively unprofitable projects.

Basic Model B.

Consider now the case where eventual innovation is assured, but its timing is uncertain. A model of this nature is considered by Dasgupta and Stiglitz (1980) and Loury (1979). Investment by each firm on a particular project undertaken at time 0, yield the desired innovation at an uncertain time. The size and nature of the innovation is fixed and independent of firms' spending. Innovation can be advanced by increased spending by each firm. Specifically, if a firm spends x at time t=0, then the probability of discovery in the interval (t,t+ t) is $\lambda(x)\Delta t)$. Where several firms compete for the innovation, assuming they pursue difference research avenues (i.e. there is no deliberate replication), the probability of discovery at any specific time is:

(30) $$\sum_{i=1}^{n} \lambda(x_i) \exp[-\sum_{i=1}^{n} \lambda(x_i)]t$$

and the expected time to discovery is

(31) $$[\sum_{i=1}^{n} \lambda(x_i)]^{-1} .$$

If all firms are equal, equation 30 simplifies to:

(30a) $$n \lambda \exp[-n\lambda t] .$$

Following Dasgupta and Stiglitz, we assume that the function $\lambda(x)$ takes a form similar to that of $h(x)$ in Figure 3. We assume that in competing for the patent, firms adjust their scale so as to increase the probability of being first, but ignore the effects of increased investment on the innovation date.* Thus competition takes the form of increasing expenditure (x) to increase the probability of being first.

Each firm i therefore maximizes

$$(32)\qquad \int_0^\infty \lambda(x_i)e^{-(\sum_{j=1}^{n}\lambda(x_j))t}\int_t^{t+T} he^{-rt}\,dtdt - x = \frac{\psi h}{r}\int_0^\infty \lambda(x_i)e^{-(\sum_j \lambda(x_j)+r)t}\,dt - x$$

As all firms are assumed equal $\sum_j \lambda(x_j) = n\lambda$.

First order condition for maximum of equation 32 is:

$$(33)\qquad \frac{\psi h\lambda'}{(n\lambda+r)r} = 1 .$$

Free entry leads to zero excess profits:

$$(34)\qquad \frac{\psi h\lambda}{(n\lambda+r)r} = x .$$

Combining equations (33) and (34) yields:

$$(35)\qquad \lambda' = \lambda/x$$

---

*This assumption is necessary to yield quantitative results. However, for small numbers of competitors it may be inappropriate. Ignoring this effect biases our results towards a shorter patent period. The extent of the bias is unlikely to be large, unless very significant economies of scale exist in research.

Marginal returns equal average returns to investment by any individual firm, so that firm size is fixed at x^* --the efficient scale, independent of patent life. (See Figure 3).

In terms of Model A we may define:

(36) $b(n) = \lambda n/(\lambda n+r)$.

Thus the results of equations (7) - (10) and (17) - (26) apply. Optimal ψ for small process innovation may be derived from equation 10.

(37) $\psi \rightarrow b'n/b = r/(\lambda n+r) = r/\psi y(\lambda+r)$

yielding

(38) $\psi = (r/y(\lambda+r))^{1/2}$

Where y is defined as before $y = b(1)h(x^*)/rx^* = \lambda h(x^*)/rx^*(\lambda+r)$ = the expected profitability to a single firm with no competitors and infinite patent life, operating at x^*.

Table 4 illustrates the optimal patent life and resulting number of firms for various combinations of $\lambda/(\lambda+r)$ and y. Note that $\lambda/(\lambda+r)$ may be roughly interpreted as the present value of the expectation of discovery. It is therefore roughly comparable to $1 - q$ in Table 1.

The results are very similar to those of Table 1. Low probability projects require longer patent periods. But the majority of projects require short patent periods in order to insure a small number of competitors and to reduce statistical duplication. While many basic research projects in the pharmaceutical industry are characterized by low success probabilities--the bulk of R&D expenditure is spent during the development state, where success probabilities are higher--so low patent life is indicated.

Table 4

Patent life and number of firms competition with

fixed scale small process innovation**

Model B

y	1.5			1.2			2			3			5			10		
$\lambda/(\lambda+r)$	T	ψ	n	T	ψ	n	T	ψ	n	T	ψ	n	T	ψ	n	T	ψ	n
.2	6.5	.73	1.08	9.0	.83	1.0*	5	.63	2.3	3.6	.52	3.8	2.6	.4	6	1.6	.28	10
.5	5.6	.67	1.0*	9.0	.83	1.0*	3.5	.5	1.0	2.7	.41	1.5	1.9	.32	2.2	1.4	.22	3.4
.7	5.6	.67	1.0*	9.0	.83	1.0*	3.5	.5	1.0*	2.3	.35	1.0*	1.4	.24	1.3	1.0	.17	2.0

*ψ set to appropriate value to insure one firm pursuing the innovation.

**These calculations are valid for any r.

Perhaps an indication of the relevant magnitudes may be obtained from Schwartman's (1976) study, which indicated an average cost per new chemical entity of about \$17 million per year with average time to discovery of about 4 - 5 years.

In Model B, average time to discovery is $1/\lambda n = 5$ years (taking the upper limite). As the current patent period of 17 years implies $\psi = .97 \simeq 1$, the zero profit condition may be written as

$y(\lambda+r) = n+r = .2 + .2 = .4$. Hence, optimal $\psi = (.2/.4)^{1/2} = .71$, implying an optimal patent life of 6.4 years. Note, however, that this calculation ignores the time required to complete the minimum experimentation necessary, so tends to extend the optimal patent life. If, for example minimum experimentation time is 2 years, optimal patent life falls to 4.8 years.

For product innovations the results are similar. Assuming first linear demand for the product, we may derive optimal ψ_p from equation (26) to yield:

(39) $\psi_p = 4b'n/3b = 4r/3(\lambda n+r) = 4r/3\psi_p y(\lambda+r)$ yielding

(40) $\psi_p = [4r/3y(\lambda+r)]^{1/2} = 1.15\ \psi$

Thus optimal patent life for product innovation, assuming linear demand, is only moderately larger than that which is optimal for small process innovation for most projects.

However, the results are quite different for constant elasticity demand curves. If quantity demanded $Q = AP^{-\eta}$. Monopoly price for the product is set where

(41) $P(1-1/\eta) = MC$. This is illustrated in Figure 4.

Defining MC = 1 for convenience, we have:

(42) $P = \eta/(\eta-1)$

and profit flow

(43) $\pi = (P-1)Q = P^{-\eta}/(\eta-1)$.

Zero profit may therefore be defined as:

(44) $\psi b(n)P^{-\eta} = (\eta-1)rxn$.

Defining private profitability y as before: $y \equiv P^{-\eta}b(1)/(\eta-1)rx$ yields

(45) $\psi b(n)y = nb(1)$.

Social welfare is composed of the present value of expected consumer surplus above the monopoly price (P) during the patent, plus consumer surplus above marginal cost (=1), after patent expiration.

(46) $$S = b(n)[\psi \int_{\eta/(\eta-1)}^{\infty} QdP + (1-\psi) \int_{1}^{\infty} QdP] = b(n)[\int_{\eta/(\eta-1)}^{\infty} P^{-\eta}dP + \int_{1}^{\infty} P^{-\eta}dP] =[\psi(\frac{\eta}{\eta-1})^{1-\eta} + 1-\psi] \frac{b(n)}{\psi-1}$$

Maximizing (46) subject to (45) yields

$\psi_{\eta} = b'n/b(1-A)$ where $A \equiv (\eta/\eta-1)/1-\eta$ and ψ_{η} indicates optimal ψ for product innovation with constant elasticity demand curves.

The ratio of ψ_n/ψ is shown in Table 5 as a function of the demand elasticity (η). As can be clearly seen, this ratio is generally above 1.15 indicated for the linear demand. As well, it grows considerably as demand elasticity falls.

The reason for the discrepancy can be illustrated by Figure 4 . At the point P_1Q_1 the linear and constant elasticity $\eta = 1.2$ demand curves are tangent. If optimal monopoly price is set at P_1 during patent life, the point P_1Q_1 represents the same private profitability for both demand curves. However, consumers surplus both during and after the patent is significantly greater for the constant elasticity curve. While the same applies to the $\eta = 2$ curve, the nature of the gains is different. As demand elasticity falls the gains during the patent time rise relative to the gains in the post patent period. Therefore the losses from postponing the patent fall relative to the gain during patent life as demand elasticity falls. Thus the benefits due to increased innovation due to longer patents rise relative to the cost of a longer patent, justifying longer patent lives.

The elasticity of demand for many prescription drugs is probably quite low, as evidenced by very high price cost ratios.* Implying that for many prescription drugs the optimal patent life for product innovation is quite long. Thus for example, where expected profitability $y = 2$ and $\lambda = .2$ optimal patent life for product innovation rises from 3.5 years for process innovation to 10 years or 12 years for demand elasticities of 1.5 and 1.2 respectively. For marginal low probability projects, an infinite patent life is optimal, even

---

*Optimal price/cost ratios are $\eta/\eta-1$.

Table 5

ψ_p/ψ as a function of η and

corresponding patent life

η	1.2	1.5	2	3
price/cost $(\eta/\eta\text{-}1)$	6	3	2	1.5
ψ_η/ψ	1.8	1.54	1.4	1.34
$y = 1.2 \quad \lambda = .05$	∞	∞	∞	∞
$y = 1.2 \quad \lambda = .2$	∞	∞	∞	∞
$y = 2 \quad \lambda = .05$	∞	16	11	9
$y = 2 \quad \lambda = .2$	12	10	6	6
$y = 5 \quad \lambda = .05$	6	5	4	4
$y = 5 \quad \lambda = .2$	4	3	3	3

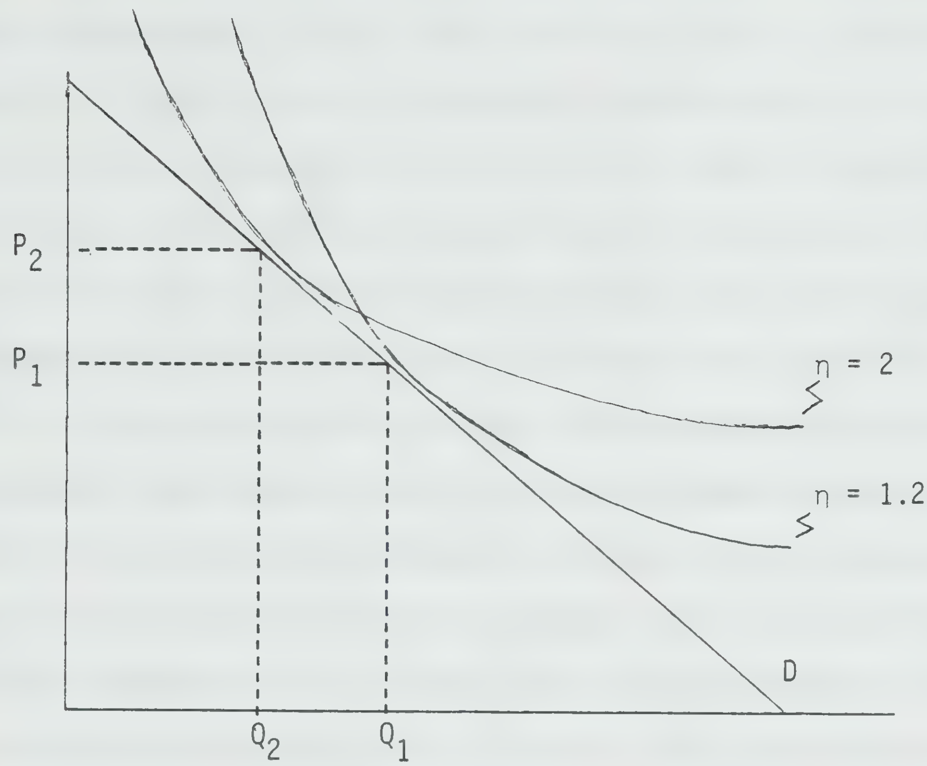


Figure 4

for higher elasticities. If net* time to discovery is 3 years, patent lives vary from 12 years for $\eta = 1.5$ to infinity for $\eta \leq 1.2$. Because the results are so sensitive to assumptions about the nature of demand curves, it is important to evaluate the nature of demand. Note, that high price cost margins do not tell us anything about the curvature of demand curves, because such low elasticities and high price/cost margins may occur in linear as well as constant elasticity demand curves. High price cost margins may simply indicate high private profitability--rather than large social benefits during the patent period.

Moreover, private demand for pharmaceuticals may be a poor indicator of social benefits. Demand elasticities may be much lower than warranted by objective considerations due to ignorance by consumers and moral hazard due to insurance coverage. This is readily apparent in the case of low demand cross elasticities among brands of the same drug. In particular, the assumption of low demand elasticities at very high prices without limit, incorporated in the constant elasticity demand curves, is unsatisfactory.

It therefore appears that the results based on the assumption of linear demand are more likely to be relevant.

Timing Considerations

Additional concavity in R&D processes may arise due to acceleration of research in response to increased competition in response to increased patent life. However, these effects are not robust. In the context of Model B, if each firm takes into account the effect of its expenditure on the overall probability of success, increased patent life increases the average efficiency of research.* However, in the context of a similar model in which R&D is continuous over time,** increased patent life leads to a reduction in average efficiency.

However, where small numbers of competitors are involved, timing considerations tend to shorten optimal patent life significantly. As most projects appear to generate relatively small numbers of competitors at optimal patent life, these considerations are important.

---

*This was not allowed for in our version. For details see Kamien & Schwartz (1982) p. 211.

**For a description of the model and proof see Kamien & Schwartz (1982) pp. 211-213. This model is similar to our Model A.

Gallini and Kotowitz (1985) have shown that competition prevents firms from following an optimal sequential research strategy, where a subset of the relevant avenues are pursued during each period. This strategy, while postponing expected discovery, saves on resources, because on the average, discovery occurs before all avenues are explored. As well learning takes place over time, so efficiency of experimentation is improved. A small number of competitors pursuing one avenue each, effectively perform research in a sequential manner. However, as numbers increase, too many avenues are pursued per unit time, reducing te efficiency of research. This may be a particularly serious problem where the number of potential avenues is small and profitability high--typically the applied research and development stages. These considerations thus tend to further shorten patent life for inframarginal projects, where short patent life is indicated by other concavity considerations.

An entirely different consideration arises for the case of marginal projects. Barzel (1964) and Kitti (1974) argue that, where markets are expanding or R&D costs are falling due to improved general knowledge, there is generally only one innovator regardless of patent life. As soon as innovation is just profitable for one innovator, someone will jump in, in order to foreclose potential rivals. Thus all innovations are marginal in this model. In order to prevent such premature innovation, it is necessary to reduce patent life.

Optimal ψ can be shown to be $\psi' = (r-s)/(r+k)$ where s is the rate of growth of potential royalties and k is the rate of decline in expenditure (x) to pursue the project. If $r = .2$ and $s = k = .05$ patent life is 11 years, and if $r = .1$ and $s = k = .03$ patent life is 6 years. These are relatively long patent lives, which arise because all projects are marginal, thus requiring a relatively long patent life.

Note that if the probability of innovation is low, as time passes without innovation--profitability increases, inducing further entry. However, as profitability rises over time, the initial patent life is no longer optimal. Higher profitability may require a shorter patent life as given in Tables 1 - 4.

This point illustrates the general deficiency of this model, which assumes a continuous improvement in R&D projects over time. However, R&D prospects do not generally proceed in this fashion. The process is more likely to be characterised by discrete jumpts in the state of knowledge, which open up a variety of R&D projects with different levels of potential profitability. Thus not all projects are marginal. The infra marginal projects require shorter patent lives in accordance with Models A and B. The model does indicate however, that patent life for marginal projects, is shortened due to considerations of technical progress and growth of demand.

It therefore appears that timing considerations are likely to contribute to shorter patent life for specific projects but may increase the proportion of marginal projects. Thus, their effect on average patent length is uncertain.

Aggregation

So far we have dealt with optimal patent life for specific projects. We have seen that optimal life is extremely sensitive to the prospective profitability of the project and to the degree of uncertainty associated with it. In general, the lower the expected profitability and the higher the degree of uncertainty--the longer the patent life which is optimal for the project. As well, product innovations generally require a longer patent life, than process innovations with equal profitability and prbabilities. Thus there is a large variance in optimal project specific patent life.

However, from a practical point of view, it is necessary to design a fairly uniform patent system in order to minimize transaction costs and an undesirable degree of government bureaucratic discretion in the administration of the patent system. Indeed, as pointed out by Wright (1983), the low informational requirements on the part of the government is the major advantage of the patent system over the alternatives. Thus, it is important to evaluate the optimal patent life for an industry or the economy as a whole. We turn now to this task.

The optimal average patent period must take into account the social benefits and costs associated with a given patent period for very different circumstances. In particular, a patent period which is too long for a given project results in a possible waste of resources and lower social benefits due to postponement of competitive use of the innovation. However, if the patent period is too short, it leads to insufficient investment in the relevant project, so that some worthwhile projects are not undertaken, or attract too few firms, reducing the probability of innovation.

It is clear that the optimal average patent period is likely to be a weighted average of the project specific ones. The precise calculations are extremely complicated and require knowledge about the nature and distribution of projects, which cannot be obtained. However, some light may be shed on the relevant considerations and orders of magnitude by aggregating projects according to some criteria and investigating the results.

We shall first assume that all projects involving small process innovations are equal in innovation probability functions $b(n)$, but differ in potential profitability (y). The frequency distribution of potential projects is defined by $f(y)$. We assume for simplicity that the costs of all projects are

the same, ($x_i = x_j$) thus differences in size of projects are embodied in the distribution f(y).* Given patent life (T) which determines $\psi \equiv 1 - e^{-rT}$, firms invest in any project until excess expected profits are eliminated in accordance with models A or B. Then scale of investment by each firm is fixed at x*. assumed to be the same for all firms and projects. The number of firms pursuing each project are determined by equation (5)

$$(5) \qquad \psi b(n)h(x^*)/rx^*n = \psi yb(n)/nb(1) = 1.$$

Social welfare is then the weighted sum of the social welfare of all projects, where the weights are the frequencies of each:

$$(47) \qquad S = (1-\psi) \int_{y_1}^{y^*} b(n)B(y)f(y)dy \ .$$

where (48) $B \equiv (h+\eta h^2/2)/r \equiv y/b(1)+r\eta y^2/2b^2(1)$ and y is as defined before ($y \equiv b(1)h(x)/rx$).

Potential profitability ranges from a minimum $y = 1$ to a maximum y^*. However, once ψ is set, some projects with low profitability may not be profitable even for one innovator. The marginal project is then defined by eauation (5) for $n = 1$. i.e.

$$(49) \qquad \psi y_1 = 1 \Rightarrow y_1 = 1/\psi \ .$$

Optimal patent life, or equivalently ψ, is found by maximizing equation (47), where equation (5) holds for each project and equation (49) defines the marginal project.

$$(50) \qquad \frac{ds}{d\psi} = \frac{1-\psi}{\psi} \int_{1/\psi}^{y^*} \frac{nb'bBf(y)dy}{b-b'n} - \int_{1/\psi}^{y^*} bBf(y)dy + \frac{(1-\psi)}{\psi^2} b(1)B(y_1)f(y_1) = 0 \ .$$

*A large project is counted as many small ones with the same y.

The first term of equation (50) defines the sum of the marginal social benefits from extending the average patent period due to increased probability of innovation in all inframarginal R&D projects, as additional firms are attracted to every project. The second term identifies the sum of the social losses due to postponement of the benefits to consumers of inframarginal projects due to increasing patent life. The third term identifies the social benefits due to the addition of marginal projects undertaken in response to increased patent life.

It is convenient to solve for optimal ψ in terms of model B. Assuming $f(y)$ uniformly distributed--i.e. $f(y) = 1$, yields a polynomial of order 4 in ψ.

For $\lambda = .2$, $y^* = 5$, $r = .2$, and $\eta = 2$, Optimal ψ is .48

and for $\lambda = .05$, $y^* = 5$, $r = .2$, and $\eta = 2$, Optimal ψ is .51.

Implying patent lives of 3.3 years and 3.6 years respectively.

The resulting marginal projects are $(y_1 = 1/\psi)y_1 \simeq 2$ for both.

These results may be compared to those of Table 4. It is immediately apparent that the inframarginal projects dominate the picture. The resulting optimal life is heavily biased in the direction of patent life appropriate for the high profitability projects. Indeed a significant tail of low profitability projects is eliminated as minimum profitability required to permit even one firm to operate exceeds 2. The reason is obvious: infra marginal projects involve large social welfare, postponement of which weighs much more heavily than the relatively low social benefits conferred by the marginal projects. Note however, that our assumption that all values of y are equally likely may have biased the results towards the inframarginal projects. It may be argued that the number of inframarginal projects is small relative to that of marginal projects, so that $f(y)$ should be skewed towards the lower end. Recall, however, that project size

was subsumed in the distribution of $f(y)$. As the development stages which involve the greater portion of R&D expenditure are likely to be more profitable than earlier stages,* the distribution of y in expenditure terms is likely to be less skewed to the left than the distribution of projects. As well, high profitability projects were truncated at a low value of 5. Even low frequencies of higher profitability projects may be sufficient to bias the results towards lower patent lives.

Another consideration involves the nature of the probabilities of discovery. As we have seen, the degree of concavity of the innovation process increases with the probability of discovery (λ)--leading to shorter patent life with a small number of innovators. The latter stages of the R&D process are characterized by higher success probabilities. As their dollar share is high, their weight will bias the average patent life towards the short end.

It therefore appears that the average patent life yielded by our experiments may be a reasonable guide to the order of magnitude involved. Even if allowance is made for increases patent life for product innovation it is unlikely that patent periods in excess of 6 years are warranted.

The preceding analysis assumed a limited degree of concavity due to statistical duplication. When actual replication is allowed, the results favour drastically shorter period of time. In terms of the model of equations (47) - (50), $b = 1$ and hence $b' = 0$. Thus the first term of equation (50) representing the gain in improved probability of inframarginal innovations vanishes. Equation 50 then becomes:

$$(51) \qquad \int_{1/\psi}^{y^*} Bf(y)dy + \frac{1-\psi}{\psi^2} B(y_1)f(y_1) = 0 \, .$$

which yields a considerably shorter patent period.

---

*This is because unprofitable projects are eliminated prior to development.

Breadth of Protection

The breadth of patent protection is an important dimension of the patent system which is particularly important for the pharmaceutical industry. There are two aspects to the problem. First there is the problem of the nature of protection. The current system protects product innovations only to the extent that they are produced by patented processes--"Product by process". Thus new products are only partly protected. As the burden of proof is on the immitator, this appears to yield fairly effective protection. However, the scope of protection is still narrow. As well, new uses for existing products are not protected--thus the incentive to investigate therapeutic properties of non-patentable products is curtailed. A related issue involves the problem of patents to imitative or competitive innovations. In the case of process innovations, where one process is a pure substitute for another, resources spent on such innovation are wasted. This arises where a competitor's innovation reduces production costs by an alternative process, which cannot be used in conjunction with an existing patented process. To the extent that such innovations arise from pre patent competition, they have been accounted for in our analysis. However, where patents enable imitators to invent around the patent at a lower cost, a patent award to the imitator is socially undesirable, because it increases

---

*Including minimum experimentation time.

private benefits to the imitator but generates no public benefits. The net social benefits of such patent protection are negative on two counts. First, the expenditure on the imitation are socially wasteful and second, the incentives to the original innovation are reduced--generating insufficient investment and reduced probability of innovation.

The situation is similar, for product innovations. Where products are very good substitutes for each other, imitation is wasteful. With very high demand elasticities between the imitation and the original product, there is very little consumer surplus due to the imitation if prices remain unchanged. There is simply a transfer of profits from the innovator to the imitator, accompanied by wasted resources on the imitation and possibly, on excessive non price competition to gain market share. As well, the prospects of such imitation reduce the incentive to the original innovations. Even if prices fall, due to competition between imitators and inovators, the reduced incentive to innovation dominates the welfare effects, provided that the original patent period was close to the optimal one.

Where imitation leads to a significantly differentiated product the situation may change. The greater the difference--the lower the demand elasticity for the imitation, so that consumer's surplus rises even during the patent period. The degree of patent protection if any, to the imitation thus depends on the degree of differentiation between an innovation and imitations. In particular, it must be related to the incremental increase in consumers surplus due to the availability of the innovation and imitation together. The incremental consumers surplus of the imitation relative to that of the innovation can be approximated by the proportional increase in the combined sales. relative to the sales of the innovator, given unchanged prices.

In any case, our analysis suggests that a relatively broad patent protection should be accorded innovators. To the extent that this is impossible, patent protection must be denied to imitators. The Canadian patent system for pharmaceuticals appears to be seriously deficient in this respect. The "product by process" patent protection is very narrow. It encourages imitation by alternative processes for products which are not differentiated. Such competition is purely wasteful.

Compulsory Licensing: Optimal Royalties

It has been pointed out (Tandon 1982) that compulsory licensing with an appropriate royalty rate, coupled with a long patent period, is always superior to a shorter patent life which yields the same present value of profit to the innovator. This is because, while supplying the same incentive to innovate, it yields greater benefits to consumers. The extent of such additional benefits is positively related to the elasticity of demand.

To see this, consider Figure 5 . Assume that a patent of one period which nets the inventor $(P-C)Q$ in profit is replaced by a two period patent with compulsory royalty of P_1-C which yields the rectangle $P_1fgc = Pabc/2$ per period. Ignoring discounting, profits to the patent holder remain unchanged. However, as output rises at p_1, $p_1-c < (P-C)/2$. With one year patent consumers surplus consists of the rectangle $Pabc$ + the triangle abd. Under licensing consumer surplus is twice the rectangle $PaeP_1$ and twice the triangle aef. However, $PaeP_1 > P_1fgc = Pabc/2$ because profits to the inventor are unchanged. As well aef > fgd. Therefore, the gains in consumer surplus in the first period ($PaeP_1$ + aef) exceed the losses in the second period (P_1fgc + fgd).

The gain arises from the fact that output rises as royalties fall. As long as social and innovator rates of discounts are equal, the preceeding logic stands. However, if innovators rate of discount is sufficiently higher than the social one, the advantage of compulsory licensing will be eliminated.

In order to investigate the order of magnitude of the combined optimal level of royalties and length of patent life, we derived both for model B. The results yield an infinite patent life for all projects with compulsory licenses. The precise optimal level of royalties as a fraction of the cost saving due to a process innovation is very difficult to calculate. However, it is always lower than the optimal value of ψ for the project. For $y = 2$, $\lambda = .2$ the royalty rate is approximately 40% of cost savings. For $y = 4$ it falls to less than 30%. For product innovations, results are more easily determined. In the case of linear demand curves optimal royalty as a fraction of innovators profits (α) is given by

$$\text{(52)} \qquad \alpha^2(1-\alpha) = r/4y(\lambda+r)$$

This yields values of α which are considerably smaller than the corresponding values of ψ. Solving for values of $y = 2$, $\lambda = .2$, $r = .2$ yields $\alpha \simeq .25$ and for values of $y = 4$, $\lambda = .2$, $r = .2$ yields $\alpha \simeq .16$. Note that contrary to the optimal patent life, optimal royalties are smaller for product than for process innovations.

While we have not solved explicitly the problem of averaging over different projects. The same considerations apply as in the calculations of optimal average patent life. Inframarginal projects dominate, leading to low average life.

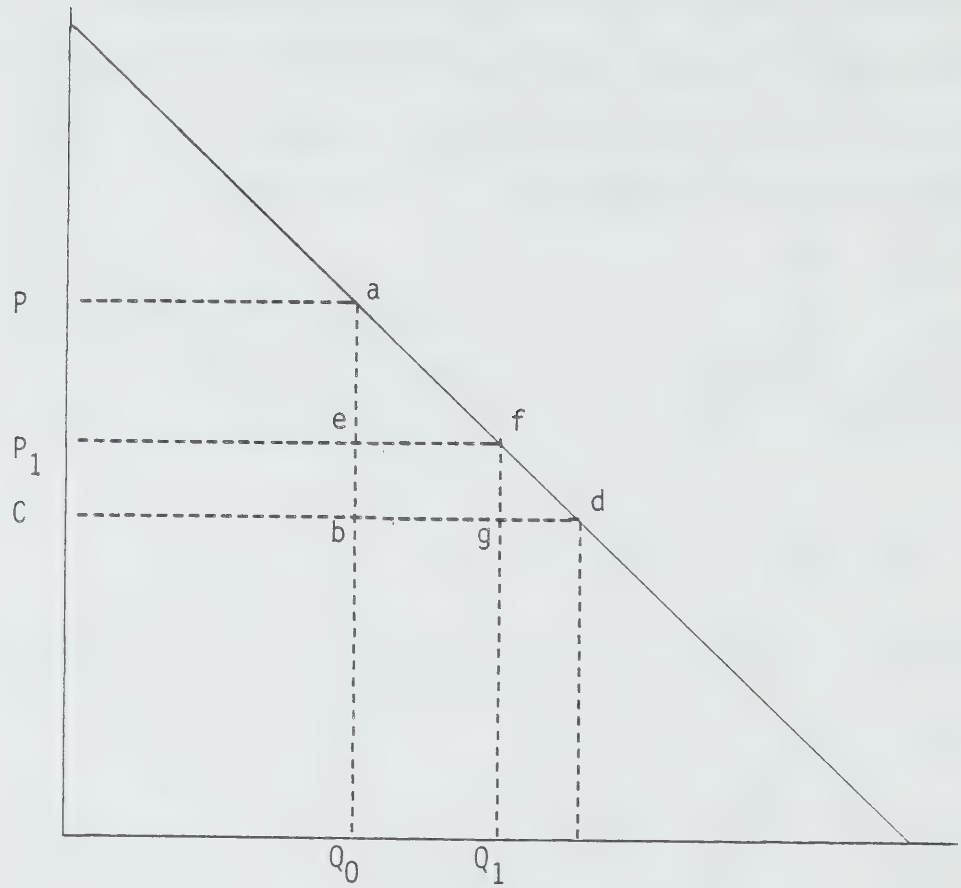


Figure 5

Optimal Patents--National Considerations

The preceding analysis utilized the social welfare function which weighted all individuals equally regardless of their position as producers, inventors or consumers and regardless of their nationality or residence. We may however, wish to evaluate patent policy from a Canadian perspective in terms of the costs and benefits to Canada rather than world-wide. The work of Berkowitz and Kotowitz (1982) suggests that because Canadians are mainly consumers of pharmaceutical products and because the market size is so small relative to world markets, the incentive effects of the Canadian patent system in terms of encouraging additional innovation are extremely small but the costs are high. As very little local production and/or commercial innovation takes place in Canada, the patent system creates a significant transfer from Canadian consumers to foreign multinationals. Even if such transfers were fully transformed into productive R&D expenditure, the benefits to Canadians from these innovations would be very small, approximately equal to our world consumption share (2%).

If the patent period in the rest of the world were optimal or shorter than optimal, it could be argued that the additional incentive to R&D created by Canadian patents, more or less offset the benefits which we get from the incentives supplied by consumers in other countries. However, if the patent system in the rest of the world is much too long, as it likely the case,increased payments by Canadians for foreign inventions bring further investment in R&D which mostly duplicates other efforts and the productivity of which is likely to be very low. Thus, while high prices paid to foreign pharmaceutical companies increase the profitsof these successful innovators, this increasing profits is largely dissipated by additional unproductive R&D conducted elsewhere in the hope of capturing some of these rents. Thus, these transfers from Canadian consumers

accrue mainly to foreign researchers but are unlikely to generate significant social benefits to anyone.

We evaluate these elements in detail and investigate the order of magnitude of optimal patent life and royalty scales utilizing the basic model's framework. Because pharmaceuticals involve mainly product innovations, we shall analyze patent protection for products, assuming linear demand. We adapt the basic models to allow a distinction between local and foreign patent periods and ignoring foreign consumers. Thus define: $\phi = 1 - e^{-rTc}$, when Tc is the length of the Canadian patent period, and c = the ratio of Canadian consumption to consumption of the rest of the world. Demand elsticity in Canada is assumed to be equal to that of the rest of the world. Patent holders are assumed to incur a fixed cost of F for expenditure related to local testing necessary to obtain approval. Present value of profits from operations in Canada (ignoring innovation costs -x) is then

(53) $\phi Gc - F \geq 0$ where $G = P_2\gamma h^2/4r$ = The present value of gross profits to the patent holder in the rest of the world under infinite patent life.*

The number of firms undertaking R&D towards any product innovation is determined on the basis of global profits, with competition reducing global profits to zero.

(54) $[\psi + \phi c]b(n)G - F = xn$ where $\psi = 1\text{-}e^{-2T}.97$ and $T = 17$ is the patent period

period in the rest of the world.**

---

*See pages 28/9 for detailed explanation and derivation.

**We ignore advertising and approval costs in the rest of the world as they are not relevant to the analysis.

Following equation (23) we may define consumer welfare in the rest of the world:

$$(55) \qquad S = b(n)G(4-3\psi)/2 \ .$$

Similarly, consumer welfare in Canada is:

$$(56) \qquad W = cb(n)G(4-3\phi)/2$$

In determining the optimal patent period in Canada, we may consider only Canadian welfare in our calculations. Alternatively we may calculte the optimal Canadian patent to maximize global social welfare. We may then compare the results.

First consider Canadian welfare:

Maximizing equation (56) subject to equations (53) and (54) we obtain:

$$(58) \qquad -3cbG+cG(4-3\phi)b'dn/d\phi = -3cbG+cG(4-3\phi)b'bnc/(b-b'n)(\psi+\phi c) = 0$$

and solving for ϕ :

$$(59) \qquad \phi = \frac{4b'nc-3(b-b'n)\psi}{3cb} \leq 0 \quad \text{as} \quad c \leq .2 \quad \text{and} \quad \psi \simeq 1, \text{ for almost any } b(n)$$

function.

Thus optimal Canadian patent period must be determined by equation (53) to insure introduction of the patent into Canada.

$$(53a) \qquad \phi \geq F/cG \ .$$

Performing similar calculations to maximize global welfare $(S+W)$ yields:

$$(60) \qquad \phi = \frac{4(1+c)b'n-3\psi b}{3cb} \simeq \frac{4}{3}\ \frac{b'n}{b} - 1 \leq 0 \quad \text{if} \quad \frac{b'n}{b} < \frac{3}{4}$$

From Tables 1 - 4, it is apparent that $\frac{b'n}{b} \leq 3/4$ for all< $y \geq 1.5$. Thus even from a global point of view, the optimal Canadian patent life must be determined solely on the basis of local consideration i.e. by equation (53a). The reason for

this, is that global patent life exceeds the optimal level considerably. Thus the marginal product of additional research effort brought forth by the small additional incentive of the Canadian patent is extremely small relative to the welfare gains to Canadians. Note that setting the patent period at $\phi = F/cG$ yield zero profit from the Canadian operation, so that from an incentive point of view ϕ F/cG is equal to $\phi = 0$.

If we allow informative advertising to affect the demand for the product in Canada, the results are somewhat different. A positive patent period is indicated even from a narrow national welfare point of view. Such advertising may be introduced by making Canadian sales a function of advertising in the following way: $c = c(a)$ where a is advertising expenditure upon the introduction of the product. We may define this advertising as informative because it does not affect the elasticity of product demand and therefore leaves price cost margins unchanged.

Profits from the Canadian operation are redefined as

$$(53b) \quad \phi c(a)G - a - F \geq 0 .$$

The level of advertising expenditure is determined by the first order condition.

$$(61) \quad \phi Gc' = 1 .$$

Thus the level of advertising is affected by the Canadian patent, introducing an additional effect of patent life on Canadian consumer welfare. Maximizing W subject to equations (53b) and (54) yields an approximate solution for ϕ .

$$(62) \quad \phi = 4/3(1+\varepsilon_a) \quad \text{or} \quad \phi = (a+F)/cG,$$

whichever is larger. Where a is determined by equation (61), and $\varepsilon_a = \frac{-cc''}{(c')^2}$ = the

degree of concavity of the advertising function $c(a)$.

We may obtain an idea of ε_a if we assume $c = \beta a^{\alpha}$. Then $\varepsilon_a = (1-\alpha)/\alpha$. Under present conditions $\phi \simeq 1$ hence equation (61) may be written as $\frac{cGc'}{c} = \frac{cG\alpha}{a} = 1$. So $\alpha = \frac{a}{cG} \simeq$ the current advertising to gross profits ratio. Note, however, that only informative advertising is relevant. It therefore appears that $.1 < \alpha < .3$, where advertising is broadly interpreted to include all selling expenditure. The resulting patent life is given by $.13 < \phi < .4$, implying $.7 < Tc < 2.6$. These are clearly shorter than the comparable globally optimal patent lives of Chapter 2.* As well they may fall short of patent lives which guarantee patentees non zero profits.

Compulsory Licensing

As we have seen in Chapter 2, long patent lives with compulsory licenses with appropriate royalties are generally superior to shorter patent lives. The argument holds equally well in the case of patent consideration for a national perspective. Thus, infinite patent lives with appropriate royalties might be desirable, with the optimal royalty rate as a proportion of gross profits $R \simeq 1/(1+\varepsilon_a)$.

Using our assumption about α yields $.1 < R < .3$.

However, further consideration suggests that a very short exclusive patent period, combined with compulsory licensing during subsequent years may be superior. There are several reasons.

---

*Chapter 2 ignored considerations of informative advertising. If these are added,globally optimal patent lives may be somewhat longer than those indicated in Chapter 2.

First, the rate of discount applied by the patentee is likely to exceed the social discount rate. This is because patent protection is imperfect, so that the private discount rate must incorporate the probability of increasing availability of substitutes over time due to non licensed imitation. This is particularly important under current patent regulations which only cover "product by process".

Second, the granting of clearance in Canada, apparently eases considerably the achievement of clearance for generic drugs in many other countries. Thus, early introduction of a drug into Canada followed by early generic entry, increases significantly the ease and speed of penetration of generic producers to other countries. However, in the absence of clearance by the patentee, generic producers or importers would find it much more expensive to obtain initial clearance. Because royalties are relatively low, postponement of clearance by a patentee threatened by immediate entry of generics, may well be desirable. Particularly when account is taken of subsequent generic entry into other markets on the coattail of Canadian approval. Thus it is necessary to allow the patentee a sufficient period in which to recoup his local regulatory and advertising costs and during which he is protected from entry by generics. Of course, royalties in subsequent periods should be correspondingly lower.

References

Barzel, Y. (1968): Optimal Timing of Innovations. Review of Economics and Statistics 50, 348-55.

Berkowitz, M. and Kotowitz, Y. (1982): Patent Policy in an Open Economy. Canadian Journal of Economics 15 (1).

C.A.A. (1983): Compulsory Licensing of Pharmaceuticals.

Dasgupta and Stiglitz (1980): Uncertainty, Industrial Structure and the Speed of R&D. Bell Journal of Economics 11, 1-28.

Eggan, Higginbotham and Weston (1983): Economics of the Pharmaceutical Industry.

Gallini, N. and Kotowitz, Y. (1985): Optimal R&D Processes and Competition, Economica 52, 321-334.

Gorecki, P. (1981): Regulating the Price of Prescription Drugs in Canada: Compulsory Licensing, Product Selection and Government Reimbursement Programs, Economic Council of Canada.

Kamien, M. I. and Schwartz, N. L., (1982): Market Structure and Innovation, Cambridge University Press.

Lefler, K. (1981): Persuasion or Information? The Economics of Prescription Drug Advertising, Journal of Law and Economics. 24(1) 45-74.

Levine, Klevorick, Nelson and Winter. (1984): Survey Research on R&D Appropriability and Technological Opportunity, Mimeo, Yale University.

Loury, G. (1979): Market Structure and Innovation, Quarterly Journal of Economics XCIII 395-410.

Nordhaus, (1969): Invention, Growth and Welfare Cambridge, Mass. M.I.T. Press.

Scherer, F. M. (1980): Industrial Market Structure and Economic Performance Chicago, Rand McNally, 2nd ed.

Schwartzman, (1976): Innovation in the Pharmaceutical Industry. Baltimore, Md., Johns Hopkins University Press.

Spence, M. (1984): Cost Reduction, Competition and Industry Performance, Econometrica, Jan.

Stiglitz, J. (1969): Discussion, American Economic Review Proceedings. May, 59, 46-49.

Tandon, P. (1982): Optimal Patents with Compulsory Licensing. Journal of Political Economy 90, 470-86.

Tandon, P. (1983): Rivalry and the Excessive Allocation of Resources to Research, Bell Journal of Economics. 14, 157-165.

Telser, L. et al. (1975): The Theory of Supply with Application to the Ethical Pharmaceutical Industry, Journal of Law and Economics 18.

Usher, I. (1964): The Welfare Economics of Invention. Economica N.S. 31. 279-87.

Wright, B. D. (1983): The Economics of Investment Incentives, American Economic Review, Sept. 691-707.

10. 8[illegible]